

GS 言語科目(教養英語科目)成績評価の 厳正化を進めるための予備研究 —成績評価の現状と教員ごとの成績評価のばらつき—

大藪 加奈
OYABU Kana

Abstract

This article is a report on the assessment of Liberal Arts compulsory English courses at a Japanese national university. The assessments examined are that of two English for Academic Purposes (EAP) courses, EAP I (Paragraph Writing) and EAP II (Public Speaking) held in the first quarter of 2016/2017 academic year. Both courses have common assessment criteria and rubric/assessment sheet. Although the overall grades of EAP I were relatively evenly distributed, grade distribution of individual classes varied substantially with a few instructors giving very generous grades. In EAP II, overall grades were more generous than EAP I, and there are also some instructors giving generous grades. By conducting interviews and analyzing questionnaires and course portfolios submitted by EAP instructors, this preliminary study identifies possible causes which may influence assessment behaviors of EAP instructors.

1. はじめに

金沢大学では、2016年度より、共通教育（教養教育）を刷新した。クォータ制を導入し、各学期の長さが8週間となったほか、「GS 言語科目」と呼ばれている1年生必修の英語科目は、EAP（English for Academic Purposes）コースとTOEIC 準備コースの授業をそれぞれ4科目ずつ、計8科目受講するカリキュラムとなった。両コースとも、共通シラバスによる英語中心の授業運営を行っている。

著者は、EAP コース教育企画部において、共通カリキュラムの作成や統一評価基準の設定を担ってきたが、小論は、新カリキュラムの同一名称科目・共通シラバス授業間における成績評価の厳正化を進めるための予備研究として、第1クォータに開講されたGS 言語科目 EAP コース2科目の成績評価の現状を調べ、教員の意識調査を行って、教員ごとの成績評価のばらつきが起こる要因を特定する試みである。

2.1. GS 言語科目の成績評価

GS 言語科目の成績評価方法は、TOEIC 準備コースでは、共通テストによる評価が全成績の80%、教員による評価が全成績の20%の割合となっている。（共通テストは、第1クォータから第3クォータまでの第8週授業時に1時間のTOEIC 形式学内模試の形で行

い、第4クオータ終了後には、TOEIC-IPを行っている。) 一方 EAP コースでは、共通の評価方法とその割合を定め、科目ごとに共通のルーブリックを使って、各担当教員が評価している。このように外部試験の活用や、EAP の観点に留意しつつ独自の到達目標を設定することは、中央教育審議会が示した外国語コミュニケーション能力の厳正な成績評価を行うための改革方策と合致していると言える。¹

2.2. GS 言語科目 EAP コースの授業科目と共通評価基準

2016 年度の GS 言語科目 EAP コースの内容は、以下のとおりである。

- ・EAP I: アカデミックライティングの基礎コース。パラグラフ (300 語) の書き方を習得する授業
- ・EAP II: パブリックスピーキングの基礎コース。一人 2 分程度 (グループで 10 分程度) の英語プレゼンテーションのしかたを学ぶ授業
- ・EAP III: EAPI と EAPII の統合型コース。要約 (100-130 語)、レスポンス・ペーパー (200 語)、比較エッセイ (300 語) を書き、英語での質疑応答のしかたなどを学ぶ授業
- ・EAP IV: EAPIII を発展させたコース。要約 (100-130 語) 3 つを含む 5 パラグラフのミニ・リサーチペーパー (500-600 語) を書き、英語での簡単なディスカッションを学ぶ授業

これらの科目には、共通評価基準を示すルーブリックや評価表が定められている。(第 1 クオータに開講された EAPI および EAPII の評価基準は Appendix 1、Appendix 2 を参照のこと。²

2.3. EAP I コース・EAP II コースの成績評価の分布

今回の調査では、2016 年度第 1 クオータに行われた EAPI と EAPII の成績分布を調べた。全クラスの成績分布は以下のとおりである。

表 1 : 2016 年度第一クオータ EAPI・EAPII 成績分布

	S	A	B	C	不可・放棄
EAP I	7%	35%	42%	11%	5%
EAP II	17%	60%	19%	2%	2%

EAP コースでは、成績分布のガイドラインを以下のように定め、EAP コース担当教員に提示している。

¹ 中央教育審議会, 「学士過程教育の構築に向けて」平成 20 年 3 月 25 日 p.26-27

² EAPIII、EAPIV の評価基準については、『外国語教育フォーラム』第 10 号 (2016 年) pp.29, 37 を参照のこと。

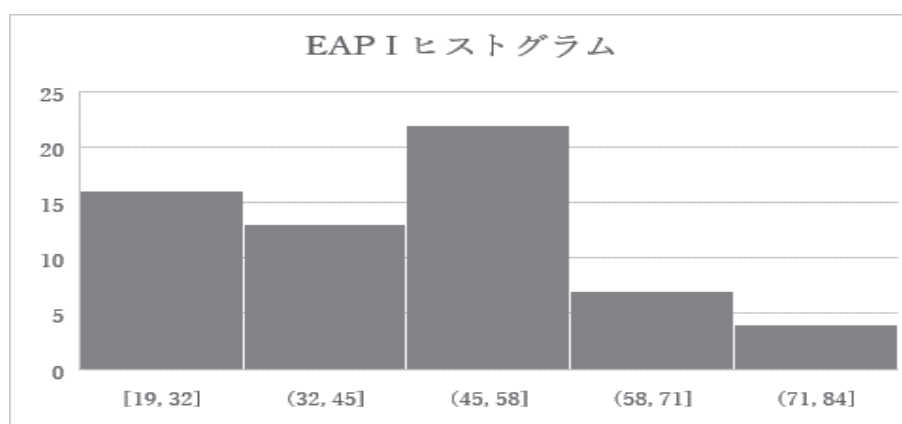
表 2：GS 言語科目 EAP コースの成績分布ガイドライン

S	A	B	C	不可・放棄
10%程度	30%程度	40%程度	20%程度	到達目標に達していない場合

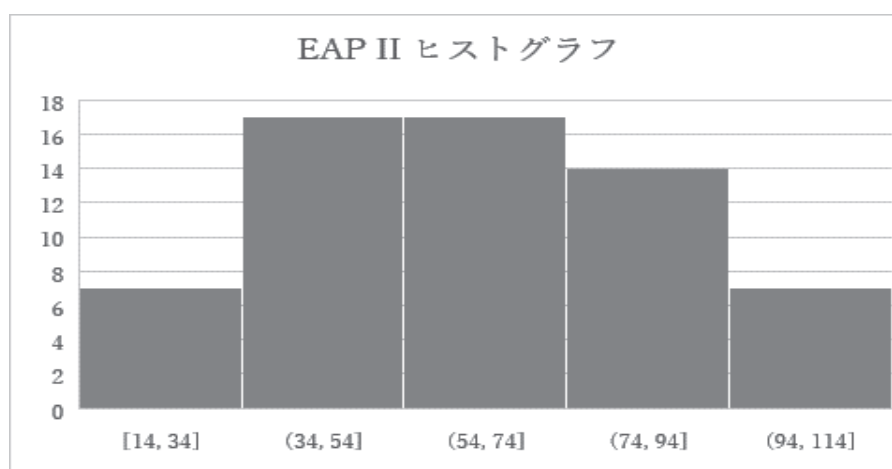
ガイドラインと実際の成績分布を比べると、EAPI は EAPII よりもガイドラインに近い分布であり、EAPII は S 評価 A 評価の割合がガイドラインよりもかなり高い。

EAPII の評価が高い原因としては、英語プレゼンテーションが初めての学生もいることを予想し、授業に出席して活動に参加し、成績評価に関わるプレゼンテーションを行った場合、60 点以上の評価となるような共通評価表を使っていることが影響している可能性がある。しかし、今回は教員の意識に焦点を当てているので、コースごとの評価分布の違いに関する分析・考察は、行わない。

グラフ 1：EAP I の S 評価と A 評価の合計の割合



グラフ 2：EAP II の S 評価と A 評価の合計の割合



2.4. クラスごとの成績評価の分布

全クラスの成績分布は上記のようになっていて、EAPI では全体として、ガイドラインに近い形になっているが、クラスごとの成績分布を見てみると、成績のつけ方にばらつきがあることがわかった。以下にそれぞれのコースにおける S 評価と A 評価の合計の割合をヒストグラムで示す。

EAPI コースでは、S 評価と A 評価の合計が、45%から 58%のクラスが最も多く、22 クラス（全体の 35%）で、ガイドライン（S 評価 10%、A 評価 30%の合計 40%）を超えた評価となっている。一方、合計の割合が 19%から 32%であったクラスも 16 ある。（全体の 26%）71%以上の高い割合で S と A の成績がつけられたクラスは 4 クラス（全体の 6.5%）あり、最も S と A の割合の合計が高いクラスでは、クラス全体の 84%が A 評価以上となっていた。

EAPII では S 評価と A 評価の合計が 34%から 54%のクラスと 54%から 74%のクラスが共に 17 クラスで最も多く、90%以上となっていたクラスは 11 あり、そのうちの 3 クラスは 100%となっていた。

2.5. 成績分布のばらつきによる問題点

GS 言語科目は、クラス割当制を取っており、学生はコンピュータでランダムにクラスに振り分けられていて、授業時間帯や担当教員を選ぶことはできない。もしクラス間の授業や評価に、担当教員の評価方法に由来する差が出る場合には、外部試験や独自の評価基準以外にも方策を講じる必要がある。旧カリキュラムでは、学生が各時間帯 2-3 クラス開講されている授業から自分が履修したい教員の授業を選ぶことができたが、それでも教員間の宿題の量などに関する不満が授業アンケートに書かれることがあった。クラス割当制になると、クラス間、教員間の差は以前よりも大きな問題となり得る。TOEIC 準備コースでは、教員が関わる評価の割合は 20%なので、教員ごとの評価の違いが成績全体に及ぼす影響は比較的小さいと言えるかもしれないが、EAP コースでは教員の評価が 100%成績に反映されるので、教員ごとに評価のしかたが違くと受講生の間に不公平感が生まれる可能性がある。また、課程選択に GPA を使っている学類では、実際に不公平となる可能性もある。

3.1. 教員間のばらつき

今回調査した EAPI と EAPII のクラスごとの S 評価と A 評価の成績分布では、S 評価と A 評価の合計がクラス全体に占める割合が少ないクラスにおいては、全体的にやや厳しめの教員はいても、特定の教員のクラスで常に割合が低いということはなかった。しかし、S 評価と A 評価の割合が非常に高いクラスについては、特定の教員が担当している比率が高いことが分かった。たとえば、S 評価と A 評価の和の割合が 90%以上の 18 クラスのうち、10 クラスは、2 人の教員が 5 クラスずつ担当しており、これらの教員は、割合が 90%を下るクラスをほとんど担当していなかった。（のこりの 8 クラスは、8 人の教員で担当していた。）この 2 名の教員は、担当学類の違いにかかわらず、一貫して S 評価と A 評価をクラス全体のほとんどの学生に与えたことになる。EAP コースは習熟度別クラスではな

いので、これらのクラスに優秀な学生が偏ったとは考えにくい。もちろん、これらのクラスの担当教員の教授法が他の教員に秀でていたために到達目標を達成できた可能性はあるので、今後更なる調査が必要である。実際、オンラインで行われている学生による授業評価アンケートでは、上記の教員は、学生に授業内容を高く評価されている。ただ、S 評価 A 評価の割合が高い教員の成績評価が、他の教員に比べて単に寛大であることも考えられるので、今後は更に厳正な成績評価を可能にする方策を講じるとともに、当面「寛大な」成績評価をする教員のクラスを特定の学生が複数回受講することで、ほかの学生より成績評価が極端に有利になることがないようにクラスわりあてのしかたが計画されている。

3.2. 教員の意識

今年度から始まった EAP 新カリキュラムでは、共通シラバスの説明シンポジウムや研修会の開催、冊子体での内容周知、共通ルーブリックや到達目標の作成、成績評価方法の統一と成績評価の割合の統一、成績評価分布ガイドラインの周知など、様々な方法で成績評価について教員へ通知してきた。それでも上記のようなばらつきが出ることについては、早急にその理由の解明と、必要に応じた対応をする必要があるが、その準備段階として、まず教員がどのように成績評価について考えているのか、教員の意識を調べることも必要であろう。今年度は、EAP 担当教員へのコース内容フィードバックシート（35 名中 12 名が回答）、EAP 研修会に参加した教員への EAP コース・ポートフォリオ（35 名中 20 名が作成）そして 4 名の EAP 担当教員への聞き取り調査を行った。

聞き取り調査・フィードバック調査、ポートフォリオ調査の結果、EAPI の成績評価で S 評価と A 評価の合計の割合が高い（80%以上）クラスを担当していた教員（High Percentage Instructor 以下 HPI）、S 評価・A 評価の割合がガイドラインに近い、またはそれよりもやや高いクラスを担当していた教員（Normal Percentage Instructor 以下 NPI）、S 評価・A 評価の割合が低い（20%以下）クラスを担当していた教員（Low Percentage Instructor 以下 LPI）の言葉を以下に拾う。

「EAPI の成績評価基準に従えば、中間評価や期末試験に加えて宿題点や授業貢献点がよほどよい学生でなければ S 評価にならない。加えて、数名の中間評価を提出しなかった学生や期末試験を受けなかった学生は -40 点からの評価になり、ほぼまじめに授業に参加していても不可にせざるを得ない場合がある。」（LPI-1）

「EAPII の評価基準に従えば、中間試験や期末試験の日に欠席していない限り、不可になることはない。スプレッドシートでつけたが、結果として成績評価にかかわるプレゼンテーションで指定時間それなりに話せたら A 評価、やや難ありで B 評価、宿題等をちゃんと出していない、授業中発言がないと C 評価となるが、S 評価はなかなかつけにくい。」（LPI-2）

「私が学んだ国の大学では、70 パーセント以上の成績がつくことはほとんどない。A 評価・S 評価はかなりよい評価という印象がある。」（NPI-1）

「私が学んだ国の大学では、Bは良くない評価であるから、学生にBをつけるのはかわいそうな気がする。」 (HPI-1)

「金沢大学の学生は非常に優秀だ。作文もよく書けているので、評価も高くなる。」 (HPI-2)

「S評価はそのクラスで断トツの学生、という位置づけだと思う。30人クラスで1～2名くらい。多くて3名いるかいらないか。」 (NPI-2)

「授業貢献点、宿題点、小テストの結果など、名簿に細かく記録して、スプレッドシートで管理している。一番よく貢献した学生、宿題等の成績の一番良い学生を100%として平常点の成績をつけている。2-3時間の宿題を課せとあるが、宿題を課すことと、その宿題がちゃんとなされているかチェックして評価することは、別物であり、すべて評価に反映させるには教員に多大な時間とエネルギーが必要だ。」 (LPI-3)

「スプレッドシートで計算し、ガイドラインの成績分布とかなり違う割合になっている時には、もう一度成績のつけ方の見直しを行っている。ガイドラインがあることで、成績の分布は良くなっていると思う。」 (NPI-3)

「教室の活動は学生全員で行っているので、授業貢献点は出席している学生ならほぼ100%になる。」 (HPI-3)

「しっかり授業をすれば、学生は到達目標をクリアして、A評価以上になると思う。B以下の評価をつけるのは、自分の教え方が良くなかったと言っているような気がする。」 (HPI-4)

「到達目標の達成を目指して、作文等の添削を丁寧に行っている。学生がよい成績を取るかどうかは、先生の添削努力に関係するのでは。時間をかけて添削しているので、学生がその内容を習得してくれていたら嬉しいのだが、、、。」 (NPI-4)

「共通ルーブリックは大まかすぎて、あまり参考にならない」 (NPI-5)

「正規分布にとらわれず、パラグラフライティングやエッセイライティングなど、決まりを習得できなければ今後英語で作文する時に困るようなアカデミックライティングを教える科目では、ほとんどの学生が目標を達成するJ型分布にすることも考えるべきだ。」 (NPI-6)

3.3. 考察

教員の学生時代の経験と成績評価

NPI-1 や HPI-1 から、教員の成績評価の考えに影響を及ぼす要因の一つとして、「自分が受けた教育」があると思われる。自分が大学教育を受けた国での状況や経験を基準にして、S 評価、A 評価、B 評価、C 評価のイメージを描いている場合、違う国・違う大学で教育を受けた同僚教員との間で、評価のしかたに違いが生じる可能性がある。HPI-4 のように、B 以下の評価を指導者の力量不足とみる教員の場合、悪い評価をつけにくいことも考えられる。

非常勤講師の本務校と本校の学生の学力差と成績評価

別の要因としては、普段教えている学生の英語力を基準にすることが、特に非常勤講師などの場合に考えられる。本務校の学生と金沢大学の学生の間に英語力の差がある場合、HPI-I のように、金沢大学生全体への評価が高くなる可能性がある。

ガイドラインの提示と成績評価

NPI-3 のことばのように、正規分布に近い成績分布のガイドラインを提示することで、教員の成績評価を客観的に見直す機会を提供し、極端に偏った成績分布を減らす可能性がある。この点に関しては、ガイドライン提示前後の成績分布を調査分析し、ガイドラインの有効性を確かめたい。

ルーブリックの質と成績評価

さらに、NPI-5 が指摘するように、教員に提示されたルーブリックは設定がおおまかすぎて、成績評価の厳正化に寄与していない、また主観的な評価となることを妨げられない、という問題もある。今年度は新カリキュラム初年度ということもあり、教員の新しい授業へのストレスを最小限にする使いやすいルーブリックを作成することを目指したが、そのためかえって多くの教員の意見が反映できるおおまかな記述となり、統一評価基準のメリットが生かせないルーブリックとなったかもしれない。実際、著者は提示されたルーブリックでは学生に評価内容が伝わりにくいので、ワークシートの一環としてさらに詳しいルーブリックを作成し、それを使って学生に自分の作文を採点させたり、推敲する際の参考にさせたりした。

教員が成績評価にかけられるエネルギーと時間

LPI-1 や HPI-3 から、統一の評価方法やその割合を決めても、たとえば宿題や授業貢献をどう評価するか、たとえば宿題の出来不出来、授業貢献の回数や度合いまで詳しく評価する必要があるのかどうか、によって評価結果は変わってくる。また、NPI-4 のように、教員がどの程度作文の添削をするのか、という問題は、教員がかけられる時間やエネルギーの差によるところも大きく、実際非常勤講師からは、作文のための添削が必要なような EAP の授業は教員の負担が非常に高いので、時給で雇われていて時間の余裕もない非常勤講師にはむずかしい、という意見もでている。

正規分布とJ分布

NPI-6が指摘するように、EAPに焦点を当てたクラスについては、正規分布よりも、ほとんどの学生が到達目標を達成できるようなシラバスや成績評価の基準を作成することも考慮する必要があるかもしれない。

4. 結び

本論では、2016年度第1クオータのEAPコースEAPI科目とEAPII科目の成績分布の現状と、教員の成績評価に関する意識調査の結果を報告した。EAPIでは、コース全体としてはガイドラインに近い成績分布になっているが、特にS評価とA評価の全体に占める割合に注目してクラスごとに比較すると、担当教員によって成績評価分布にばらつきがあることが、わかった。また、EAPIIクラスでは、成績評価基準の影響もあってか、全体的に高い評価になっているが、ここでもS評価とA評価の全体に占める割合は、担当教員によってばらつきがあることが、わかった。聞き取り調査やフィードバックシートの調査、コース・ポートフォリオの記述の調査から、教員の成績評価に影響を与える可能性のある要因もいくつか特定できた。新カリキュラムは始まったばかりであるが、今後成績評価に関する研究をさらに進め、その結果を生かした統一シラバス・統一評価基準の改革を行い、より厳正で透明な成績評価を達成できるような科目運営を進めたい。

Appendix

EAP I Assessment Criteria for Mid-term Assignment and Final Exam

S: 90-100%

- Well-organized and coherent paragraph with convincing reasons and examples
- Just a few grammatical and spelling mistakes
- Correct paragraph format, properly using indent, margins on both sides, double-space, Times New Roman font, and 12 point font size

A: 80-89%

- Coherent paragraph with sufficient reasons and examples
- Some grammatical and spelling mistakes
- Correct paragraph format, properly using indent, margins on both sides, double-space, Times New Roman font, and 12 point font size

B: 70-79%

- Satisfactory paragraph which includes the three main components but insufficient logical development of ideas
- Several grammatical and spelling mistakes
- Incomplete paragraph format, including 1 or 2 minor mistakes in relation to indent, margins on both sides, double-space, Times New Roman font, and 12 point font size

C: 60-69%

- Incomplete paragraph which includes disorganized paragraph components with insufficient word lengths
- Many grammatical and spelling mistakes
- Incomplete paragraph format, including a few minor mistakes in relation to indent, margins on both sides, double-space, Times New Roman font, and 12 point font size

F: 0-59%

- Disorganized and incoherent paragraph with one or two missing paragraph components and insufficient word length
- Too many grammatical and spelling mistakes
- Wrong paragraph format with insufficient word length

Public Speaking Assessment Sheet

Student name:

Group:

	Good (5)	Average(4)	Poor (3)
<u>Organization</u>			
Introduction			
Transition			
Conclusion			
<u>Content</u>			
Appropriate thesis/ main point			
Evidence and explanation to support main points			
Consideration of audience			
Appropriate length			
<u>Fluency and Delivery</u>			
Eye contact			
Voice projection/ tone			
Gesture/ movement/ posture			
Pace/ Use of pause			
Pronunciation/ Intonation			
Use of visuals			
Use of notes			
<u>Language</u>			
Vocabulary			
Grammar			

Overall grade:

Criteria highlighted in grey are assessed per group; those in white are assessed per student.

The overall percentage is reached by dividing the total score by 80 then multiplying by 100. (Total score/ 80 X 100)